

Data sharing: A viable resource for future

Kritarth Naman M. Singh, Yashashri C. Shetty

Department of Pharmacology and Therapeutics, Seth G.S. Medical College and K.E.M. Hospital, Mumbai, Maharashtra, India

Abstract

Clinical trials and research studies are being conducted worldwide at a rampant pace leading to generation of large amount of data. However, to reap the benefits of the data generated it is important that this data is shared with the general public without which it can be deemed useless. Despite its importance being known to us, data sharing does not come without its share of problems and it is not as easy to execute as it sounds on-paper. Over the past few years, multiple coveted organizations around the world involved in research activities have come up with their respective guidelines and initiatives to make sure the sharing of research data is smooth and ethical. Developing countries like India have made a few strides in the right direction with some initiatives in-place, but there still seems a long way to go before unanimous data sharing can be a reality. The stakeholders may have to face certain possible repercussions due to data sharing but there is no doubt that if done in the right way, it can lead to universal development.

Keywords: Clinical trial, data sharing, guidelines, India, research

Address for correspondence:

Dr. Kritarth Naman M. Singh, Department of Pharmacology and Therapeutics, Seth G.S. Medical College and K.E.M. Hospital, Parel, Mumbai - 400 012, Maharashtra, India. E-mail: kritarth1989@gmail.com

INTRODUCTION

Clinical trial, as defined by Schedule “Y” of the Drugs and Cosmetics Act (2005), intends to find answers for a valid research question, which is executed after collecting data from the study population, analyzing the data and disproving a hypothesis.^[1] The late 1970s and 1980s saw a wide recognition of the importance of trial data which was statistically more reliable, leading to systematic reviews as well as meta-analysis. This also led to the formation of various international organizations with a view to collaborate on researches based on different but similar clinical trials.^[2] As time passed, the number of research studies being conducted has grown multi-fold,

which is obviously leading to the generation of huge amount of data. It is obvious that if we want to reap the rewards of the huge amount of data being developed, it should be shared so that the data is interpreted as well as reused for future research. However, the process of sharing data is complex and is based on various factors which include the type of research and the public policy. In many cases, there is no expertise on the part of the researchers to share the data while sometimes; the data may not be transferable. Ethical concern is one of the major hindrances in transferring of trial data.^[3] This article looks at various aspects of clinical trial data sharing, with its importance as well as the recommendations proposed by various institutions worldwide.

Access this article online

Quick Response Code:



Website:

www.picronline.org

DOI:

10.4103/2229-3485.203036

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

How to cite this article: Singh KM, Shetty YC. Data sharing: A viable resource for future. *Perspect Clin Res* 2017;8:63-7.

IMPORTANCE OF CLINICAL TRIAL DATA SHARING

Sharing of clinical trial data has gained momentum in the last few years after the researchers worldwide started to gauge its importance. Data can be present in various forms – either in the form of “raw data” found in case record forms of trials or in “coded form” which is present in computer databases. However, data sharing holds true to any kind of data and usually done after the study is completed. It is mainly executed by publishing the study results in a scientific journal. However, sharing of data before publication may help in situations when the immediate benefit to the society is necessary (for instance, if a trial drug was found to cause serious adverse events to participants). Data sharing helps in deducing results of single clinical trials, and one can compile the data of multiple studies to get to a strong conclusion.

There have been instances where the participants of clinical trials have died after receiving the test drug. Recently, in January 2016, at least one such participant died in France, and four other participants were injured after receiving a fatty acid amide hydrolase inhibitor, under-trial in humans for the first time.^[4] Although such fatal instances are rare, it is important that such data be released from the manufacturer at the earliest so that further damage to human life can be prevented.^[5] Data sharing helps in assessing the risk-to-benefit ratio of various interventions and thus, will reduce the exposure of trial participants to deleterious effects of an intervention which has already been tested before. It is possible to detect any inaccuracies in previous research studies, and this will help in awakening the need for a reanalysis of the data or the intervention. Data sharing leads to the development of a new research question and helps in the acceleration of research. Overall, data sharing helps in restoration of trust in clinical research and promotes the growth of science.^[6,7]

PROBLEMS ASSOCIATED WITH DATA SHARING: WHOSE DATA IT IS??

The need and importance of sharing clinical trial data are associated with a plethora of barriers. The investigators are concerned that they put months or years in collecting their clinical trial data, so they must be given some time to analyze the data and communicate their findings with public.^[8] However, it is a well-known fact that only a section of the trials conducted get their results published quickly. A recent finding was that <50% of the studies funded by the National Institute of Health (NIH) are published within 2½ years of study completion.^[9]

Data sharing is a complicated process, and for having an efficient data sharing system, it is important that the study

teams have collected the data accurately and in a well-defined way. Data preservation may not be a priority in many situations, especially when the resources are limited. Many a times, the language of the data set may pose a hindrance in understanding and interpreting the data. Compatibility of the data sets with modern software systems may prove to be another roadblock.^[6]

The ownership of the data set is an important question which has been raised every now and then. There is a general feeling that the data sets are owned by the sponsor, however the study report cannot be approved without a final review by the investigators. The physicians and the patients are involved in the study too, so they should also have a partial ownership of the clinical trial.^[10] Another way of seeing this problem is understanding the ownership of a data set which may be subsequently used by other researchers for future publications. Trust is important between the data provider and the user of the data set, essential to rule out any intentional misuse of the data.

Data sharing should consider protection of participant confidentiality. The tools used for de-identification of the study population, like statistical data masking, may not be available every time.^[11] There have always been ethical concerns about data sharing, and researchers have been confused on the inclusion of a “data sharing clause” in the informed consent document itself.^[12]

Data sharing has the potential to help the progression of science, but it requires human and technical expertise. The preparation of proper data sets, the technical equipment with good speed internet connection plus people who are well-versed with the procedures are needed for an efficient process of data sharing.^[6]

There has been a universal increase in the use of electronic case report forms (eCRF) for collecting data in research. The United States Food and Drug Administration and the European-Union have published regulatory requirements for the electronic trial data systems, which should be followed for the electronic-based data collection. The information attributability and the system integrity should be maintained at all times during the study. The system has been accepted worldwide mainly due to the advantage of obtaining data from multiple sites around the world and accessing it from a single system, just with a click on the computer. This helps in easier preparation of data sets without much hassle of storing the paper documents. Not much has been said or published about sharing of eCRF data, but it is obvious that such a data set is easier to share with a second party. However, it is important that the researchers who are collecting the data enter the same on the online database in the right manner and have the necessary technical expertise to make the eCRF data credible.^[13,14]

INITIATIVES AND GUIDELINES

The importance of sharing clinical trial data has been identified by researchers and primary federal agencies worldwide. Few of the agencies have given their own guidance.^[15]

The Food and Drug Administration Modernization Act of 1997 had made it compulsory for all the trials involving drugs being tested for serious clinical conditions to be registered on an online database known as ClinicalTrials.gov. This was followed by an expansion in 2002 wherein the controlled clinical trials after phase 2 will also be included. After a demand from the researchers for access to raw data of the clinical trials, many pharmaceutical companies like Roche and GlaxoSmithKline adopted policies allowing access to data of trials for approved products.^[15,16]

In June 2013, the European Medicines Agency (EMA) made an announcement that the agency will make available some of the clinical trial data present in market dossiers. The announcement also involved sharing of certain data and study reports. However, the EMA has declared that they would not allow the access to data which they think will endanger the privacy of the participants. The researchers who have requested for the data need to assure that it will be used only to address the issue of public health and in accordance with the purposes for which the participants gave consent.^[17]

NIH has considered all types of data for sharing. In the view of NIH, "Data should be made as widely and freely available as possible while safeguarding the privacy of participants, and protecting confidential and proprietary data. Beginning from October 2003, it is expected from the researchers who are requesting \$500,000 or more of direct costs to include a plan to share the collected research data, and in case it cannot be shared to state the reason."^[18]

Cancer Research UK (CRUK), world's leading charity funder towards cancer research, has its stance on data sharing. The guidelines released by CRUK mention neither the timing nor the intricate procedures of data sharing but mention the principles for the same. Applicable since April 2009, the guidelines mention that all the applicants seeking funds from CRUK must submit a data sharing plan, which will be reviewed along with the funding decision. The guidelines mention the aspects to be included in the plan for sharing data, the content, standards, metadata, method, timescale, preservation plan, and restrictions (if any).^[19]

Organization for Economic Co-operation and Development (OECD), which addresses the economic, social and environmental challenges of globalization, has developed its

guidelines to facilitate cost-effective ways to grant public access of research data. The guideline states - "The value of data lies in their use. Full and open access to scientific data should be adopted as the international norm for the exchange of scientific data derived from publicly funded research."^[20]

Wellcome Trust, the world's largest medical research charity funder for human and animal research, has mentioned about the data sharing policy in its publication requirements for research projects funded by them. It mentions that they expect the research team to increase the access of data for the general public, with the minimum possible restrictions. The guideline also mentions that the data should be submitted in the recognized data storage systems, and in places where such systems do not exist, new resources like Dryad or FigShare can be used. If the data are held by the investigator, then an active process should be described in the publications by a contact detail or a weblink.^[21]

The Japan Science and Technology Agency has a National Bioscience Database Center (NBDC), which is a location where human research-related data can be shared. The NBDC human database works on three main principles:

- Maximum collection of human data produced with the public funds
- Sharing the maximum of this collected data with as many users possible
- Managing this data in the best way without breaching the privacy of the individual participants.

NBDC has classified data sets into mainly four types:

- Open data - Data openly available without restrictions
- Controlled-Access data - Data present for investigators with research experience who can use this data for their studies
- Data for future release - Data which is planned to be released after the results on it are published
- De-identified data - Intended to be used internally only.

The NBDC data sharing guidelines mention the general responsibilities of the researchers and the participants in relation to the data to be shared. It states that the informed consent document for the research participants must have a mention of the data sharing policy. The user should inform the NBDC office about the data usage each year in August. Once the data usage is done, the user should delete the data and inform the same to the NBDC Human Data Review Board Office.^[22]

The International Committee of Medical Journal Editors (ICMJE) has launched a proposal for promoting responsible sharing of the research data.^[23] Earlier this year

in 2016, the ICMJE proposed that the authors must provide the individual-patient data within 6 months of publication. The data sharing plan should be a part of the clinical trial registration process with details of data housing, mechanism of data access and other elements of data sharing.^[24] ICMJE has also recommended that trial registries should follow the footsteps of ClinicalTrials.gov and include mechanisms for including data sharing plans. The proposal also mentions the need to take care of the rights of the investigators and trial sponsors along with those of the trial participants. To protect the rights of the trial participants, the ICMJE have recommended the following:

- Authors who are using shared data for secondary analyses should give an attested statement that it was in accordance with the agreed terms
- A reference should be provided to give proper credit to the data generators
- Authors should comment on how their analyses differ from the primary data analysis.
- Restricted data (accessible only through process of registration and authorization)
- Sensitive data (defined in various acts and rules of the Government of India)
- Sharable data (nonsensitive in nature)
- Negative list (nonsharable data as declared by the departments/organizations).

The objective of this policy was to promote access of sharable data owned by the Indian Government through an Indian network within a predefined framework so that a wide access and usage of public data can be permitted. The usage of common standards for collecting, transferring and integrating data sets will be more feasible. The owners should be given due credit for the collection and development of data. Decision-making tasks which include plan development, asset management, environment protection, and improvement of living conditions can be executed with easy access to important data. The data transfer should be an open policy for better accessibility to its users. The policy also mentions about the online analytical processing capabilities and that the database should be multi-dimensional and oriented to a particular subject.

India has limited guidelines in place for an efficient sharing of clinical trials data. However, the Indian Ministry of Science and Technology as well as a governmental organization National AIDS Control Organization (NACO) have come up with their guidelines to address data sharing.

According to the NDSAP, the data will remain a property of the primary researcher, and if needed, the price will be decided by the owner of the data and government policy. In its draft policy, it was mentioned that the concerned ministry or department will upload minimum five high-value sets of data on “data (dot) gov (dot) in” within 3 months of this policy notification.^[26]

The guidelines classify the different types of sharable data in categories, namely, data that can be shared without a review, data that can be shared after review or national level data requiring approval from NACO. Individual level data of beneficiaries cannot be shared, thus in this regard, no proposal is allowed. NACO states certain important protocols to be followed for sharing data. The applicant should provide the data request in the proper format with defined purpose and details about the required data along with an undertaking to assure NACO that confidentiality will be maintained, conflict of interest will be avoided, and NACO will be acknowledged in the publication. If a student requires data for his thesis from reporting centers, he can collect the data, subject to ethical clearance and permissions. All the data request will be examined by NACO, and the Data Sharing Committee's decision will be based on whether the topic for analysis is in line with the priorities, presence of any conflicts of interest, and the availability of all the necessary clearances. Personal information will be filtered to conceal individual identities. It is clearly stated that the data will be shared only after approval from the secretary of NACO.^[25]

Indian Council of Medical Research, the apex body in India for formulating and promotion of biomedical research, has drafted the revised version of its ethical guidelines titled “National Ethical Guidelines for Biomedical and Health Research involving Human Participants, 2016.” The new draft states the need for managing the informational risk associated with data sharing. It mentions that raw data sets are currently being shared worldwide and special care should be taken to maintain the confidentiality of participants, location, and the research setting. When the researcher publishes the results, it is expected that all information about the research (including the final data) will be available in public domain. The new draft mentions the necessity to de-identify the data unless required otherwise for which permissions should be taken.^[27]

DATA SHARING: REPERCUSSIONS ON THE STAKEHOLDERS

National Data Sharing and Accessibility Policy (NDSAP) was released in 2012 by the Indian Ministry of Science and Technology. The policy describes various types of data which include:

The importance of taking a trial participant's consent for data sharing before the initiation of study may help in protecting his rights. However, the participants may become

apprehensive about the data sharing clause in the informed consent document, doubting the credibility of de-identification procedures. It is the duty of the researchers and the sponsor to make sure that the parties receiving the primary data respect the confidentiality of individual participants.

On the other hand, the researchers may think that though they are the ones who are involved in making the study protocol, getting Ethics Committee permission, collecting and analyzing the data, the primary findings will be shared with a party, who have done no work for the research. The researchers may be apprehensive about receiving the appreciation or acknowledgment for their hard work after sharing their hard-earned data.

Many a times, the data which are collected are of sensitive nature, discussion over which can evoke strong emotional responses from the public. If such kinds of data sets are open for sharing, they can lead to stigmatization or fear and possible political threats to the research team. Hence, such kind of sensitive data should be handled with care and Ethics Committees should act as guardians to such processes.^[28]

CONCLUSION

Data sharing is the key for universal development if privacy and confidentiality of data are maintained.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

- Schedule Y. Requirements and Guidelines for Permission to Import and/or Manufacture of New Drugs for Sale or to Undertake Clinical Trials; 2005. Available from: http://www.cdsco.nic.in/html/D&C_Rules_Schedule_Y.pdf. [Last accessed on 2016 Jun 15].
- Chalmers I, Altman DG, McHaffie H, Owens N, Cooke RW. Data sharing among data monitoring committees and responsibilities to patients and science. *Trials* 2013;14:102.
- Borgman C. The conundrum of sharing research data. *J Am Soc Inf Sci Technol* 2012;63:1059-78.
- Butler D, Callaway E. Scientists in the dark after French clinical trial proves fatal. *Nature* 2016;529:263-4.
- Improve Early Access to Data from Catastrophic Clinical Trials: A Statement on Behalf of the British Pharmacological Society. *Society News*; 2016. Available from: <https://www.bps.ac.uk/news-events/news/society-news/articles/improve-early-access-to-data-from-catastrophic-clinical-trials>. [Last accessed on 2016 Jun 16].
- van Panhuis WG, Paul P, Emerson C, Grefenstette J, Wilder R, Herbst AJ, et al. A systematic review of barriers to data sharing in public health. *BMC Public Health* 2014;14:1144.
- Institute of Medicine (US). *Sharing Clinical Research Data: Workshop Summary*. Washington, (DC): National Academies Press (US); 29 March, 2013. 2, the Benefits of Data Sharing. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK137823/>. [Last accessed on 2016 Jun 18].
- Ross JS, Lehman R, Gross CP. The importance of clinical trial data sharing: Toward more open science. *Circ Cardiovasc Qual Outcomes* 2012;5:238-40.
- Ross JS, Tse T, Zarin DA, Xu H, Zhou L, Krumholz HM. Publication of NIH funded trials registered in ClinicalTrials.gov: Cross sectional analysis. *BMJ* 2012;344:d7292.
- Suvarna VR. Sharing individual patient data from clinical trials. *Perspect Clin Res* 2015;6:71-2.
- Anderson M, Seltzer W. Federal statistical confidentiality and business data: Twentieth century challenges and continuing issues. *J Priv Confid* 2009;1:7-52.
- Pearce N, Smith AH. Data sharing: Not as simple as it seems. *Environ Health* 2011;10:107.
- Ene-Iordache B, Carminati S, Antiga L, Rubis N, Ruggerenti P, Remuzzi G, et al. Developing regulatory-compliant electronic case report forms for clinical trials: Experience with the demand trial. *J Am Med Inform Assoc* 2009;16:404-8.
- Connell CA. ERES Compliance and Data Integrity Issues in e-Trials. 2003. Available from: <https://www.socra.org/publications/past-socra-source-articles/eres-compliance-and-data-integrity-issues-in-e-trials/>. [Last accessed on 2016 Jun 12].
- Hamel M, Mello M, Francer J, Wilenzick M, Teden P, Bierer B, et al. Preparing for responsible sharing of clinical trial data. *N Engl J Med* 2013;369:1651-8.
- Idem. Principles on Conduct of Clinical Trials and Communication of Clinical Trial Results; 2009. Available from: http://www.phrma.org/sites/default/files/pdf/042009_clinical_trial_principles_final_0.pdf. [Last accessed on 2016 Jun 21].
- Watson R. European medicines agency changes policy on clinical trial data publication. *BMJ* 2014;348:4073.
- NIH Data Sharing Policy and Implementation Guidance; March 5, 2003. Available from: https://www.grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm. [Last accessed on 2016 Jun 22].
- Data Sharing Guidelines | Cancer Research UK. Available from: http://www.cancerresearchuk.org/sites/default/files/data_sharing_policy.pdf. [Last accessed on 2016 Jun 25].
- OECD Principles and Guidelines for Access to Research Data from Public Funding; 2007. Available from: <https://www.oecd.org/sti/sci-tech/38500813.pdf>. [Last accessed on 2016 Jun 23].
- Requirements for Publishing Wellcome Trust-funded Research Papers. Available from: <https://www.wellcome.ac.uk/sites/default/files/wtp053977.pdf>. [Last accessed on 2016 Jun 24].
- NBDC Human Data Sharing Guidelines. Ver. 2.0; 2015. Available from: <http://www.humandbs.biosciencedbc.jp/en/guidelines/data-sharing-guidelines>. [Last accessed on 2016 Jun 25].
- Taichman DB, Backus J, Baethge C, Baucher H, de Leeuw PW, Drazen JM, et al. Sharing clinical trial data: A proposal from the International Committee of Medical Journal Editors. *Chin Med J (Engl)* 2016;129:127-8.
- Institute of Medicine. *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk*. Washington, DC: National Academies Press; 2015.
- Data Sharing Guidelines: National AIDS Control Organization; 2015. Available from: http://www.naco.gov.in/upload/2015%20MSLNS/Data%20Sharing%20Guidelines_%2013%20July%202015.pdf. [Last accessed on 2016 Jul 01].
- National Data Sharing and Accessibility Policy (NDSAP); 2012. Available from: <http://www.ogpl.gov.in/NDSAP/NDSAP-30Jan2012.pdf>. [Last accessed on 2016 Jul 05].
- Draft: National Ethical Guidelines for Biomedical and Health Research Involving Human Participants. ICMR; 2016. Available from: http://www.icmr.nic.in/icmrnews/draft_ethical/Draft%20National%20Ethical%20Guidelines%20for%20Biomedical%20and%20Health%20Research%20Involving%20Human%20Participants,%202016.pdf. [Last accessed on 2016 Sep 01].
- McCosker H, Barnard A, Gerber R. Undertaking sensitive research: Issues and strategies for meeting the safety needs of all participants. *Forum Qual Soc Res* 2001;2:1-10.